Data underlying research paper "Exploring potential contributions of open data intermediaries" by Ashraf Shaharudin, Bastiaan van Loenen, and Marijn Janssen from Delft University of Technology (TU Delft), the Netherlands

This file is the item #4 of the folder that contains data underlying the research paper "Exploring potential contributions of open data intermediaries" (working title). It consists of:

1. Tentative interview questions (semi-structured interview)
2. Informed consent form template (verbal interview & written interview)
3. De-identified interview transcripts
4. **Coding results**

Note about the de-identified interview transcripts (and coding results):
We removed personally identifiable information from the transcripts. A few interviewees may risk being identifiable if their organisation is known. Hence, we removed the identification of the organisation and country in all transcripts. Partially disclosing the organisation or country for some transcripts increases the risks of identifying the non-disclosed transcripts.
With verbal communication, some sentences may be less incomprehensible in writing, especially if English is the speaker's second language. Thus, we did minimal edits when transcribing to improve the comprehensibility where necessary, but the main objective was to keep the transcript as close to verbatim as possible.
All interviewees whose interview transcripts are recorded in this document give permission for the anonymised transcript of their interview, with personally identifiable information redacted, to be shared in 4TU.ResearchData repository so it can be used for future research and learning.

Acknowledgement:

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 955569. The opinions expressed in this document reflect only the author's view and in no way reflect the European Commission's opinions. The European Commission is not responsible for any use that may be made of the information it contains.

Quotation Content	Codes
Reference discoverability is also a nice thing. I think for a long time, and also the INSPIRE directive states you should describe your metadata according to ISO standards. ISO metadata is brilliant in describing spatial metadata but basically it's non existent for generic administrative open data, they're working with DKAN and all kind of other formats.	C01: Different data/metadata standards
For us, it's more the issue of licensing. One issue that now turns out is open data is not open data. For instance, we are trying to combine governmental data with data from OpenStreetMap -- and different license. We have a license on governmental data which is very open, it's similar to Creative Commons. While for Openstreetmap, you will might know, it's an ODBL, the share alike and you cannot combine open data from a very free license to share like. So open does not be open -- it's open but it's different. And once we talk open data being the key to use and share data very easily, we now find out that with different licensing in open data community itself, we do not really overcome that issue at once.	C02: Different open data licenses
But the only limit for <redacted> is that it's a geoportal, it's all about geo information. And of course there's a lot of other types of information as well and <redacted> doesn't provide it. So there, you have to make a connection for people who want combinations of geoinformation and administrative information or all kinds of other useful information they need	C03: Siloed open data domains
I think another aspect which I think should be very much in focus when we speak about, well, whatever we call them open data ecosystem or whatever, is the aspect of the domain. I mean, we speak about a spatial data ecosystem, but what we see is that – and we have years and many years of building up an infrastructure, spatial data infrastructure –, but what we can see is that much of the rapid development that is actually creating value right now, is happening outside the domain. So, I think the biggest risk is that we keep on being a bit siloed, and not well connected. Still being the supplier instead of being well connected part of the development going on in other domains.	C03: Siloed open data domains
the technological drivers are embedded in other domains	C03: Siloed open data domains
The speaking about one open data ecosystem, I don't think that really mirrors the reality. I think there is a lot of soft data ecosystems and the linking together of these will be, not only a driver, but it would also be a need in order to provide these more complex needs and well, to save our planet, right? So, the more complex the problems get that we need to solve, the more we need to collaborate across these many open data ecosystem that exists	C03: Siloed open data domains
And as we made, we thought that the spatial approach to data was very bad known inside the organization. So we started to teach or our colleagues in other areas and <redacted> to use this kind of information and to share our point of view and how the best way to approach data when you don't have data available, it is just pointing in a map and you get what's available.	C03: Siloed open data domains
keep it as simple as possible because geospatial data, it goes back to land surveyors, and they're really focusing accuracy -- centimetre, millimetre -- so it's always -- there's, this drive, they're driving for even more accurate, more reliable data, almost striving for perfection. And if you want to serve as many users as possible, and also if you want to serve use cases that are relevant from a societal point of view -- for instance in <redacted>, we have this housing problem, we have nitrogen problem, climate problem, energy problem -- if you want to be able to use open data in that field it will -- because basically everyone says that they're all spatial problems, where you want to build, you don't have room for nature or to create solar farms, so basically we don't that much space in <redacted>, so it should be data-driven approaches. But I think still not enough people realise that it will only happen if the data is so simple to use that climate specialists or agriculture specialists can use it. As long as we say, oh, we have great data, but you should ask us to help you, then it's never going to run. So the whole idea of having successful infrastructure from a user perspective is to lowering the threshold to use it as much as possible. And I think going to the way of web standards, more generic standards. Spatial is not special anymore. That's also really valid for technology and for standards and how you interact with the data.	C03: Siloed open data domains C04: High technical threshold for non-specialist user
Of course, when we're talking about open data we want on one side, we've worked on open government data that the government is opening data as open data. But the open data ecosystems not yet on that level that the economy is publishing data at our open data portal in the big mass. The science for example, they say we cannot work with this open data portal, because our data is in a different way, we work on Zenodo, we publish our open data information or transcripts or diploma work. So in the future we have to work more on that the economy is publishing open data, not only profiting from our open government data.	C03: Siloed open data domains C10: Limited open data from non-govt sectors
Evolution of existing open data infrastructures aimed at the creation of sectorial and interrelated data spaces.	C03: Siloed open data domains P01: Implement federated architecture
There are already a lot of data usually user, so instead of saying we should have more data users that understand this geospatial niches, no, we should make it simpler. Accept the fact that the majority of users is not geospatial expert. They're just interested in data that might or might have no spatial component. And it's just an attribute. So you should deal with that	C03: Siloed open data domains P02: Integrate data

I think the shift that's currently topic of debate is offering services that integrate data. We have a lot of different data sources, we have different <redacted> or even more data sources of course that do not have this formal status but basically they are still organized almost like independent silos. Addresses optimize for addresses alone. Large scale topography with large scale topography. But the user is not interested -- there is not single user: oh I want to access the address registry or I want topography. I want to know if they have information about buildings and maybe if you want to know everything about this building, some of it is starting the address registries, large scale topography, the small scale typography, the real estate, tax data set. So basically to serve this user you have to almost play the role as the intermediary to know where to collect all the tiny bits of information and combine it to really come up with an answer because the current open data infrastructure is basically do it your own infrastructure. You have a problem? OK, you can search for data. We can point you to different data set. You can make a connection or download the data which you have to do the analysis yourself. We give you the building blocks for solving your problem, which you still have to solve your problem with yourself. So basically you should be capable of using QGIS or Esri software to combine this data to do your analysis and to get your answer. And the data specialist is perfectly fine with it, but a climate specialist, he could not be bothered to use that kind of system. They want to know how many charging station for electrical vehicles are in city center. And he doesn't care if the data is coming from different data sources. He just should be able to ask the question and get an answer.	C03: Siloed open data domains P02: Integrate data
You have to be specialist in order to understand the data, to understand the services involved in accessing the data. So basically it's quite a high threshold for reuse of open data. You need to be specialists	C04: High technical threshold for non-specialist user
The last, the need of focusing attention on users and the use cases, the problems to be solved rather than focusing on the data provider. Sometimes we forget the use cases and the user and their needs -- and I think that's an error.	C05: Unfulfilled users' needs
They are mostly from professional users, so to say so, rarely from citizens, but that can be really small companies like independent … how do you call them in in English? The people who are a business owner so to say, one man business. But also large companies and also all kinds of government users because the government users, they are obliged to report back if they encounter a failure. There are lots of lots of reports come from the professional users within the government.	C06: Limited feedback from lay users
For the states, <redacted> said something about the financial situation, it's a point.	C07: Loss of open data providers' revenue
for the agencies in the <redacted>, it may be several millions; for them it's an issue.	C07: Loss of open data providers' revenue
we have to go further into ecosystem thinking and also the end users have to go into a more equal system thinking whereas of now, they are just harvesting a lot of value free of charge. So, they're not really giving anything back. So, we have to develop the cooperation further. So it's more also in the mind of the private sector to give back to the ecosystem and maybe in the future also like really help supply or update data in new models of doing business or cooperating together across public and private sector. But I think if we get stuck in this situation where we give this infrastructure with the last value, just free of charge with no obligation to give anything back, I don't find that very sustainable in the long run, we have to develop it further than that.	C08: Limited value return from data reuse
So legal is really a valid point, but also they're struggling with funding because open data is a success and if you make calculations about added value for open data, it's always a positive business case. But the place where the benefits go are different places than places where the costs are being made.	C08: Limited value return from data reuse
we want to work on more visibility of the reuse of data. We publish data as open data but we don't get in return what this person did with this data? How they used it? Because some civil workers in the administration, they would like to know in certain time what happened with the data that they do a visualization, that they do a platform. This is also good for them to motivate them. But also to see what has accomplished out of it, what they can use for their work. Maybe there is something developed that they can be used for their work, so there is a more problem of giving and getting in return concerning open data.	C09: Limited use cases visibility
to get a better awareness of open data and the positive effects to convince those who do not open data to do it, not by force of law, but voluntarily	C09: Limited use cases visibility
And also to communicate to the sector in the simplest way possible the requirements to be met and highlighting the benefits for society, the adoption of all the open data, and not only for society, but also the benefits for the organization itself.	C09: Limited use cases visibility C16: Limited knowledge of open data providers
I don't think so, because sometimes I know that -- I think the telecommunications and energy companies have very good geographical systems, geographical information systems, and they don't share this data. Maybe in interchange of information or some way they could give it to the public, but I don't think they [do it]. Maybe they would sell it to other companies.I don't know. Maybe it's there, I don't know the case, but I don't think they give their information as open data. I don't know.	C10: Limited open data from non-govt sectors

And they can generate with a mix of governmental data and the economy data for big projects. We have for example, from the economy side, we have a project with one company which is doing the registrations of companies. We do the registrations of companies because everyone who's doing a hairdresser salon or even a shisha bar has to register in the government their name and where they open this business. And there are a lot of economies like from the startup company, the coworking space like <redacted> or <redacted> where they collect data from the new startups, which are accomplished. And when we put all this information together, we have the companies and the businesses in the streets, but we also have these startups which are working in this coworking space and suddenly we have picture of <redacted> economical potential which shows us more information. So these are a little projects that we are doing. But of course in the field of renewable energy, we could profit from the private solar panels where implemented on the rooftops. We could profit from the companies who are also using our solar potential map. They're profiting from our solar potential map, but we would also like to know from these companies where did they implement solar panels or how is their business developed so far. So there are some information that we want to get from the economy but we don't have this open data law for the ecosystems. The law is just for us, for the government to do. But we would also like to have cooperation with the economy where we could say if we have a certain order -- how to say -- a contract with a company, for instance, we can put in a contract a few sentences, like clause, open data clause, where we can decide if you're working on this project, we want to get data which is generated in this project as open data in machine readable versions as Excel format or CSV format so that we can publish this data as open data in an open data portal. But this is only possible when the government is working with private companies, or companies in general, on a project or when we're funding a project for them, so we can say as a return, we want some data from you. But in general, we don't have a law yet to say these companies have to do open data, or have to deliver once a year, this kind of data because the government needs this information. This is where we are not there yet. But for the civil agents for example they can also do open data. We have <redacted> where everyone can get an account and if they want to do open data or collecting data on some project, they could also publish it as open data. So this is also possible for companies and also for civil people to use our open data portal on a publishing side, not only using	C10: Limited open data from non-govt sectors
And it's also the important that data and services from the private sector should be integrated with the public resources to span the number of problems that can be solved -- for telecommunication companies, energy companies, banks, etcetera.	C10: Limited open data from non-govt sectors P02: Integrate data
But at the moment, only the government is responsible for open data and this civil tech community is the one who's desiring a good open data work. But on the other hand, they could support us. The government has to do their work, of course. And we look at what can we do for open data, but we also have to look what job we do beside open data. So it is too much pressure on the government at the moment and the ecosystem itself. The economy wants for data from the government for their business models, the civic community wants to know what happens in the government for transparency reasons, goals of open data. But we need to learn to work in this ecosystem together. So this is where I'm also asking the economy not only to wait for money from the government to work on something, but to deliver open data and do projects with us.	C10: Limited open data from non-govt sectors P15: Promote open non-govt data
I understand that the ecosystems, they put a challenge and they just try to find out how the community organizes itself. And that to me is a challenge how will, in principle, the institutions, the governmental institutions, be able to organize themselves without binding law and without perhaps some supervision body that tells them how to do	C11: Lack incentives for publishing open data voluntarily
If it's a voluntarily action, and there is no guy, no law forcing you, you need to have benefits and I'm not sure that this works in <redacted> at least.	C11: Lack incentives for publishing open data voluntarily
And there, you can't force any of those companies to do it, because of course they have to find a way to earn money with that as well so that customers have to pay.	C12: Requiring viable business model
I think the realisation is really down that that we need all these data sources in order to solve these actual factual problems. Now, for example, if you want to have a look at mobility, the charging stations for electric cars, they are currently not in the base register because at that moment when it was created <redacted> years ago, you had phone booths on the streets, but no electrical cars. There are already quite a number of parties have the state-level available, but if you want to make policies on municipal level one, do we have enough of these charging stations to make sure that we create more sustainable mobility? But when you're looking on your own data, we do not know. But the data is there. So I think it's really important that the next step of this ecosystem, private parties are not only allowed as a user, but also as a data source. So that's important lesson	C13: Overlooked non-govt open data

And I think the last importantly from infrastructure point of view is that we could basically say governments that are running these kind of infrastructures, they take the initiative and of course they design it and companies participate in the ecosystem. But basically they're quite strict dividing between public parties, they come up with the design, they make the rules; and private parties as long as they play according to the rules, they can be part of the ecosystem, but they do not have a real influence in the way in how the infrastructure is going to develop. And the problem with that the governments are, in general, not the most adaptive organizations. I think private parties are much more capable in adapting to changing circumstances, to changes into technology. So there should be -- I think you cannot create completely equal playing field, but it should be made more effort than now. We should acknowledge the strong points of these private parties. Because I think the old way of thinking within government, which, for instance in <redacted> we have the <redacted> large scale topography and I think of all the trees that are in <redacted>, only about 4% of the trees is registered in this open data set. Now if you want to do an analysis on heat, heat islands within urban environments, basically, if you have only 4% of the trees, you can't say anything about the impact of having more or less trees on these heat islands within the city. There are at least two private initiatives as run by small combination of Lidar and all kinds of data, they created data sets of basically 99.5 or 99.9% of the trees. So the data is already there. I think the whole reaction of our government, oh, we lack data, we should collect it. Whereas, especially in this Federated system, OK, other parties have already have it. They are willing to make it available. Just make clear how they can be part of this Federated system and then some of the data sources are public, some of the data sources are private, maybe some data sources are released in some kind of public private partnership.	C13: Overlooked non-govt open data C14: Practical constraints in multistakeholder engagement
We can always be better in touch with each other. And we are open to it, but also to a certain limit. We can't talk every week to every software provider or intermediary or whatever. So we organize our user community a little bit on regular meetings and there they are welcome and then they can give their feedback. And lots of those feedback is very welcome because then we can improve our products and our services. And sometimes when they keep asking, do me this format or give me that part of the information which we don't provide as we're also limited by law to what we provide, we don't provide every everything. But to the extent what we allowed to then we try to fulfill the wishes, within the limitations of possibilities and budgets, of course	C14: Practical constraints in multistakeholder engagement
If I may add, you touched upon something that I think is a general weak spot. And you can actually see it in our name: <redacted>, and this has for many years been some kind of strategic guideline. That you supply, you have easy access, you have well documented interfaces, but this about actually going in and being a consultant for the private sector, to help them understand how to use the data. We have not done, because this has been a task that was lying in the private sector, but what we can see now lately within the last couple of years is that there is a huge task here that is not lifted from the private sector. So, this is the reason why we are changing from data supplier to data facilitator. The whole ecosystem thinking that we actually need, as <redacted> say, to be much more user driven, much more out there on the other side instead of just supplying, we actually need to facilitate the uptake of data. <Another interviewer:> Yeah, and the same movement as you speak upon within the geographical ecosystem is also undergoing within the <redacted>. And within this strategic framework plan, there are some focus areas around building partnerships with the private sector and establishing ecosystem. So that's the strategic goal on that.	C14: Practical constraints in multistakeholder engagement
Well, there is, If you really want a sustainable ecosystem, you need to engage also with the private sector, not only their valuable data, but also their take and they're being closer to the end user. And in order to do that they need to be some kind of "what's in it for me", for the private sector. And as I see it right now, at least in <redacted>, there is a lack here. I mean, it's coming, it's growing but very slow and it is difficult up front to demonstrate why should they engage in this ecosystem, I mean, they just take the data and they make business on it and all is good. Why should they feed into this? So, we need to solve that in order to really get it sustainable.	C14: Practical constraints in multistakeholder engagement
Then there are some mess with too much regulation and different formats. And so some organizations, even the organization doesn't know the proper regulation to apply in every case. And that happened, for example, in the case of metadata, which there is a European regulation directive. Sometimes they are difficult to implement for some local level organization who don't have some means to implement that regulation	C16: Limited knowledge of open data providers
Development of training actions aimed not only at learning how to work with data, but also at understanding the leading role of data in today's society, as well as how to challenge existing power imbalances around data in order to better adapt them to public interests	C16: Limited knowledge of open data providers
Well with I think that -- first we are dealing with that -- people see that when you are talking with this kind of stuff, the georeferencing, they understand that there's another task that you are putting into the management of the data. And what we understand is that data since the very first start must be georeferenced, so when you manage it in your app, you are given an added value, but the position is from birth. So you don't have to do anything that you aren't doing now. But the benefits you get after is greater because data is georeferenced from the start. So we are trying to involve our areas in understanding that the since the very first moment they start a task, instead of writing a next row in a database, they think in a position in the map. And they start to grow the business with that position, because with that position, they can access much more information than just a row on a database.	C16: Limited knowledge of open data providers C21: Technical difficulties in establishing data management systems

But what I see – what I really meant about the data knowledge – is that I see the ICT domain It's really developing a lot of interesting, also ecosystems you could say, and cloud to edge and whatever technology and it's super fine systems. But when you then need to add the content, the data, it doesn't fit because they didn't have any data knowledge on board. So, it's quite important that the ecosystems dealing with data are providing that data knowledge very early on to the other sectors, in order in order to get the real value.	C17: Limited knowledge of open data users
Some users or organization, don't know how to use interoperable formats such for example GML. I don't know. Maybe you don't know it because it's very specific. Or GeoPackage which are open formats. And other downloads services. This lack of knowledge forces us to publish geographic information sometimes in non-interoperable formats because we are very committed to offer in open format, but user, non-specialized user sometimes doesn't know what to do with that kind of information	C17: Limited knowledge of open data users P04: Customise data
But it does cost something and we have to get the information and we have to make products and we have to service it and the platform like <redacted> costs quite a lot of money too. So as long as there's enough budget to keep that running, then it's very sustainable. I mean, but yeah, it depends on the budget.	C19: Incurring maintenance costs to providers
there are costs involved in open data. Somebody has to pay for them. At the moment we get our budget for this kind of open data information from <redacted>. So they provide the budget. But society and the ministry and <inaudible>, they always want more. But the budget is usually don't grow accordingly. Sometimes, there is budget for developments as well, but most of the times that's limited. So that hinders you in the development sometimes. And if you have for instance paid products, then you can talk to the people who want the developments and they can start paying for it. And in this case with open data, it's not possible	C20: Incurring development costs to providers
Another thing to take in account is that as users get used to have more and more information, they demand the information to have great quality, great updating. The technology goes so fast and they asked to have it more quickly and it is the cost for us to be up to date in the newest technology and progressing day by day.	C20: Incurring development costs to providers
And of course, like any other area, you have to maintain it and develop it further to meet the future needs. So, currently we are seeking extra funding on the finance law in order to ensure the further development so that the value does not disintegrate overtime.	C20: Incurring development costs to providers
If you really want to have a successful digital government, you need really serious funding. it needs to be on the political agenda as well. For instance, if you have a look at <redacted>, everything about data and digital transformation is a political issue. Their Prime Minister has an IT background. <redacted> So as a result, there is a completely different mindset within government. And as a result, <redacted> government is developing in a ridiculous pace, if you compared to <redacted> pace because they have the political support, the political backup, and they have the funds. So they are working with solids and with data poles and in all kinds of technical solutions to make the old data network more ethical again and make it work with regard to privacy as basically try dealing with these issues. In <redacted>, we see that we have a problem, but nobody's building a solution yet because of that issues.	C20: Incurring development costs to providers C25: Inflexible/unclear government-market boundary C26: Reliance on political agenda
That I think was one of the problems of INSPIRE, a lot of little requirements and not so many organizations have the money and the time enough to implement it.	C21: Technical difficulties in establishing data management systems
Implement processes to improve the management, quality and governance of open data in the context of the overall data strategy of each administration	C21: Technical difficulties in establishing data management systems
the reality is that a lot of the municipalities and governmental bodies that have to work with the data, they are not capable of creating everything from scratch. They don't have the funds, they don't have the staff that's capable of creating a local infrastructure on their own. So they rely on basically, these full service providers like <redacted> and other competitors. They can say, OK, we know that your task is this, this and this, we can help you with it. We take care of data storage, we take care of visualisation, we take care of how you combine all those, all these different data sources. So I think <redacted> also played an important role in lowering the thresholds of using data.	C21: Technical difficulties in establishing data management systems
Once we provide the information we have to give a step more, allowing to taking advantage of the big data based technologies and allow to create automatic processes through artificial intelligence and using clouds as technological support.	C21: Technical difficulties in establishing data management systems P05: Offer process automation
What did not work was the idea to have a uniform data model, so the data specifications they turned out to be too complicated and too inflexible. It's one issue that the Commission now tackles with the revision of INSPIRE. So that environment that we created was not really succeeding in harmonizing the data. That's probably a challenge that will be forwarded to the data ecosystems. So harmonizing the data is probably to be done with the user in focus and the user, every user, has different application. So there not be one way to harmonize it, there will be several ways. There need to be somebody who does it, and that person, the institution that does it, needs to be funded. That needs to be an interest in something like one.	C22: Complex/rigid data standards

For instance, <redacted> in the in the beginning, there's some quite nice examples, there was a map with the quality of surface water if you want to go swimming. In the summer it was first day of 25 plus degrees, and one of the news I think at <inaudible> starts with an item, there's now these websites, it's <inaudible>, so you can check your local swimming spots, you know the water is actually healthy or not. Within 5 minutes, they had two million users. So you need to be scalable. And then it pops out, OK, some of these standards, they work nice with one user, with five users, with 1000 users, maybe 100,000 users, but 2 million, it's not going to scale anymore. Basically if you want to have a real sound and grown up ecosystem, you have to be able to do that. You have to be able to scale, to deal with these kind of big requests. And I think new generation of standards does it, and so being model.	C22: Complex/rigid data standards
And the same thing I think from really detail point of view, core additional requirements are valid. I mean there are not nonsense but basically, their mindset was: let's think of the most complex use case that we want to serve with this infrastructure. OK, what requirements does the infrastructure has to take in order to be able to serve this most complex use case. And as a result, there's a lot of complexity through the infrastructure that's only really needed in the most complex use cases. But in other 90% of the use case they are much less complex. They are quite straightforward and even if you want to do something that's quite straightforward, you're confronted with all this unnecessary complexity. So basically they added all the complexity for just the few rare occasions that also the use case is really complex. What we also mean with this adaptive and agile ecosystem is that it should match up, for instance, we see that within OGC standards nowadays being modularized. Otherwise, the standard was huge and basically it has so many requirements in order to show off all use cases. Now the new OGC standards to have <inaudible> as the core, just a small set of requirements. And with a small set of regulations you can serve say 80% of the use cases. If you want to support more functions case, you maybe need to implement an extension or two extensions. So basically then there's more of a balance. If your use case is more complex, your implementation will become more complex as well because you have to implement more of these extensions or modules in order to also serve some more complex use cases. But the nice thing is that as long as you have simple use case, it's enough to work on with the core.	C22: Complex/rigid data standards
fostering inter-administrative collaboration to generate data exchanges and facilitate their openness, identifying some datasets to work on their quality and on the use of standards to really be able to obtain all the value they provide.	C23: Heterogeneous data administration
big challenge is a <redacted: type of administrative governance> in <redacted>, I think you notice it. And another point is the license, we talked about it. I think these are the two big challenges.	C23: Heterogeneous data administration
For instance, my personal experience is that I am in negotiations with <redacted> and <redacted> is a key customer of data and with the [European] Commission and they of course, would like to profit from the data ecosystem. However, they have clearly told us they would do not want to deal with <redacted: sub-national administrative levels> in <redacted>. They want one contact. With a data ecosystem that we put on the situation as it is, they would have to deal with <redacted>. They do not really like to do, but what is the solution? Does <redacted> needs to adapt? Should <redacted> stepped? Or will there be a third party, say <redacted>, who does the job and provides -- takes the data from <redacted> and provide it to European Commission. That's probably are not what the <redacted> leaders would like. So for me ecosystem still a challenge.	C23: Heterogeneous data administration
At the moment we have situation that approximately half of the <redacted: sub-national levels> has open data, others have not. The perspective is that, with HVD from June next year, in theory, any of the states would have to. However, we notice that some of the <redacted: sub-national levels> try to escape and they find gaps. Such gaps are, for instance, data privacy issues, and such gaps are, for instance, legal mandates by third parties.	C23: Heterogeneous data administration
So the situation in <redacted> is very heterogeneous and that will probably […] during the following interview. The role of <redacted>, so we started as <redacted> after World War Two, which was <redacted>, and <redacted> years ago the scope of the agency has been redefined and from that date we are definitely perceived as a provider of geodata to all the institutions of the <redacted>. So we do not necessarily produce the data, but we are intermediary ourselves. So we get the data from others, in particular from the official mapping agencies of the <redacted: sub-national administrative levels>, and we process that data, we combine it and we provide it with the <redacted> government.	C23: Heterogeneous data administration P02: Integrate data
But the only limit for <redacted> is that it's a geoportal, it's all about geo information. And of course there's a lot of other types of information as well and <redacted> doesn't provide it. So there, you have to make a connection for people who want combinations of geoinformation and administrative information or all kinds of other useful information they need. And the Ministry of <redacted>, of course, has to provide the legal regulations for combination of that kind of information together because now there's quite a lot of information available and people start combining this information and through combining you can draw more specific conclusions, you get more specific results. And which also enter into the privacy of people. And we've got maps and there's data set or open data of buildings and addresses and combining buildings and addresses and maps and aerial images and whatever you can easily get to where people live and how the environment is and whatever is possible. And that has to be regulated, of course.	C24: Privacy concern

There's also some disadvantage about open data. In terms of, for instance, our aerial image is open data, and there are companies who are providing services with that open data. We now encounter the situation that we want to do similar services for the whole country and we are not allowed to because we've got a law in Europe which says there has to be fair play between governments and companies. They make the products and services, they earn money with it, the government is not allowed to give it away for free. I don't know how, how to say it. <redacted>. And because of that law, it's not allowed to provide certain open data products although we could and we want to. Now, that's the disadvantage of open data, because there are those companies that can get the aerial image, can do their job and earn money with it, which is perfectly fine, I'm not against it, but it limits us in our possibilities	C25: Inflexible/unclear government-market boundary
The only way for us to provide that kind of products is when there is a legal ground to provide those products so that we can say it's not based on this law anymore, it's our task as a government to provide this kind of product. And so far, this list of products which is legally allowed is very limited. And together with the <redacted>, we are looking at how can we broaden this open data family, so to say, but it has to have a legal ground.	C25: Inflexible/unclear government-market boundary
we as a public sector agency might need to step a bit up on not just being a supplier, but also providing a platform that could disseminate the knowledge we have about data. And this is what I touched upon, this data knowledge. How do you actually use them? How do you connect to them? How do you integrate them? How can we make good data models? And things like that. And maybe that's lifting up the thing to a more – yeah, maybe it's some kind of data intermediaries–, but I think we need to step up from being a supplier and to be a facilitator.	C25: Inflexible/unclear government-market boundary
I think in the beginning, the <redacted> open data policy was really strict. You can only publish open data as is and as soon as you make it a little bit more adapted, a little bit more to what user needs, there's this discussion of only private parties are allowed to do that, you should not compete as a government with private parties. I think the way and how people think about what's the role of public, what's the role of private, it's also shifting a bit	C25: Inflexible/unclear government-market boundary
I think that's something that <redacted> can do much better. And the <redacted> is still a bit reluctant, is it the task of a public platform or is that something that if you want to do that, you should go to a private company and pay for it? And that's still a debate. Bit it hinders the actual uptake of data-driven approaches within all this fields that are typically not that tech oriented, or that data oriented	C25: Inflexible/unclear government-market boundary
Legal concerns is really a blocking issue. At this moment in time, <redacted> also thinking about changing law on this point and making it a specific task because <redacted> has its own law where it says what <redacted> should do. The role of having a data platform and access is already in the law, but integration of data is not yet in the law. They could change the law and they're thinking about it	C25: Inflexible/unclear government-market boundary
It's not actually mentioned in our law that strictly. And we are supposed to create value for society, but we have a long competition, and it was challenged when we freed our data in the first time by a some private sector companies. And yeah, it's always a question about drawing the line and that is why it's so important that we collaborate. And that is why it's so important that we actually expand the ecosystem because it's not one agency or private company that should do it all, we need to do it together. So, forming this new collaboration – and whether we would then become an intermediary or provider to an intermediary, I don't know, it doesn't matter –but as long as we kind of, do it together.	C25: Inflexible/unclear government-market boundary P09: Foster public-private collaboration
And then, of course, there are some more businesses like all technical aspects around the ecosystem where you have some issues. Maybe with the platforms distribution channels or maybe a large overall issue – I think – across public data is the data quality that has to be addressed. Also because we are looking now into, or seeing also in effect, that the data are being used within new areas of application not thought about when we established the <redacted>. So we have to work on metadata and the data quality so that we ensure that the quality meets the user needs, and that you do not use data for something it cannot be used for and so on and so for. Yeah, there's a large overall theme around data quality I think also.	C27: Inflexible governance/law
But changing the law takes years and years and years. Technology, it's going much faster and much higher pace. So, basically, the legal part of the framework cannot keep up with the developments on the technological point of view. So now we have a directive that basically says, you should use outdated standards, you should use complex standards. There are already lighter, easier, more generic standards available, for instance if you have look at the web feature service, the way how you can transport vector spatial data from one system to another.	C27: Inflexible governance/law
Basically, the new generation of standards is completely based on the Rest API design principles, and basically everybody that's used to working with data on web, regardless whether they are geospatial experts or not, as long as you're working with data as a developer, basically within 5 or 10 minutes you have this API up and running because it's a more predictable way of interacting with the data. So what we mean with being flexible and adaptive is that as soon as these new standards emerge, they are lowering the threshold to participate in this ecosystem. Then it's really important that you organise your infrastructure in such a way that you can adapt, that you can say: OK, we are no longer only using the old standards and it's also OK, in order to fulfill all the legal requirements, if you choose the new standards. So we have to be flexible.	C27: Inflexible governance/law

I think we now have also concepts like federated architectures that we see that it's OK to have multiple platforms, to have multiple access points for data, and maybe some access points is more from geospatial perspective with those kind of standards, and others more from administrative point of view. And it's OK that they are both there as long as they interchange as much as possible with their data. They link to each other, for instance. So there's the principles calls in some <redacted> government documents: no wrong door principle. It doesn't matter where you enter as a user, the important role of the infrastructure, it should help you regardless which door you enter. So for instance if you enter the administrative door because you don't know that there's also a geospatial door when looking for data, you should be able to discover also spatial data behind that door. And maybe, for instance, data in metadata DKAN format, where you will need really specific method that's only in ISO, it should link you to the register where the method is recorded or published in ISO standard so that you can have a look at it small additional fields that are maybe really crucial for you as a user. It's also the other way around, if somebody enters the geospatial door and said, well, but I'm looking for administrative data, you should not say, well, so we don't have data. You should still link them and help them, and basically it's not the fault of the user to take the wrong door.	P01: Implement federated architecture
I think especially the nice thing about this whole idea about federated architecture is that there is not one single central platform. And because if you think my ideal solution is 1 central platform, yeah, as long as there are two well, which one to pick? Is this one more important for the administration, is this one more important for geospatial, which one? With federated architecture, we understand, they both have valid reason to exist. They both have additional value to the users. The only thing that we should take care of is that the user is not hindered by the fact that there are two separate solutions. So they should link, they should cooperate and they should serve the user regardless their background. So I think that's an important lesson. When talking about federated architectures, the Internet is the federated architecture. It's proven it works. It works for documents and it's working for data more and more. It's one of the big steps for the Internet from the web document where Web page change to web with data where an object links to another object which links to another object. This building may link to university, but it also may link to this address, it may link to the municipality. So basically how you can navigate through pages, you can also navigate through data.	P01: Implement federated architecture
There is, I think an actual risk. At the same time, I think you can – one of the best ways is to adopt standards of the web. If they all use comparable standards in how to access data and how to discover data, use the same API patterns this time, maybe in five years time, it's something else, maybe it's linked data or it's both, I don't know. I think both approaches are valid and they will find each other because basically they are trying to do the same thing. And I think it's almost use case dependent. In the current <redacted> case, it's quite helpful, but it's coming a bit more from an architectural point of view because it solves some of the current problems that we see in <redacted> or the obstacles that we encounter in <redacted>. But I think it's European goals that they are more thematic. This is also valid point of view. Basically, they're both valid. They're more user interface let's say, more of the higher <inaudible> of why are you working on this data, on this architecture or anything -- it's a synergy between the two.	P01: Implement federated architecture
But the role I think, is essential to adapt, to transform, integrate geospatial resources, both data and services, so they can be used by society in an easy and intuitive way that maybe is a lack of organization. Maybe public organization are not so close to the final user as intermediaries are	P02: Integrate data
Yes, but not only open data, but geo open data. The ones you get the position of every data, you get more information of data itself. Maybe data by itself are some kind of abstract and with geo position, this data in a map, you see what's happening in the city. And there are problems that data are related to data. You can make a classical join like in a database with this field is joined with this. But no, the joint is spatial. You see things are happening in the same point of the city, and that's what we are thinking about.	P02: Integrate data
I think one of the drivers also from <redacted>, I think also within Europe, we are one of the front runners in that we really try to transform from only our geospatial needs into making geospatial data available through the standards of the web. Because the web is already a federated system of data, it is an ecosystem, it functions with millions and millions and millions of users. Whereas the geospatial part is just a tiny, tiny fraction of that.	P03: Transform data into open standards
And the more that we can align with those more generic standards of the web and how we describe data and how we make data discoverable, how we describe it, it would make metadata easier to get. Because I think for most use cases, if you want to work with geospatial data, you need the developer that's used to working with data. But the geospatial aspect, which is really important in the actual collection of the data, for creating the use case for the developer, I would say 90%, it's not relevant to understand how spatial works but you can create of offer kind of query or whatever kind of service you want to build. Especially to make the data ecosystem much more open and inclusive, so that it should made easy for all data developer, basically everybody who has some kind of experience with data, whether you're a journalist or a developer or a company, you should be able to work with the data. The whole idea is that by making the standards easier, by making infrastructure easier, it would lower the threshold of engaging in the system and taking advantage of the availability of open data. That's really a thing that we're standing for.	P03: Transform data into open standards

I think they could be more open to adopt, for example, INSPIRE standards, which has been very difficult and also providing greater knowledge about the needs of society to public administrations to direct our strategies to solve the real problems of citizens--that could be a good point.	P03: Transform data into open standards P19: Facilitate feedback on open data
Well, one of the things users of our open data information always ask is to give it in different formats. And we provide information in internationally recognized standard formats so that it's open to everybody. We don't do any industry specific formats or company specific formats. We just don't do it. It's not that we are not allowed to, but if we do one we have to do them all and it gets messy and troublesome and so on. So what we want is we provide it in an internationally recognized standards and we want the companies, the markets, to provide it in all the industry standards because we also recognize that our format is not always the most useful for all kinds of users, architects and building companies and whatever they want it in DXF and anything. Now we've got 3D and they want it in BIM IFC and whatever, that's fine with us, but we we're not going to do it. And there, in general, the industry and the market can do better. And <redacted>, to be honest, is a good example. For their clients, they provide a lot of information and their industry specific standards. But all the other software companies don't do that -- at least less, let me put it that way. And then architects and building companies start asking us why don't you provide it in DXF 11 and 12 and 13 and 15 and whatever. And then we say, well, we don't get the budget for that. It's not our task to do that. That's for the market	P04: Customise data
Yeah, I think <redacted> actually has said it, because this is about customizing data to a specific need or specific use case is becoming more and more important. So, I think they can play that role.	P04: Customise data
They play a role and it's an important role because in terms of, as an example <redacted>, they provide to all their customers -- they get all the open data and they provide them in a format which is for their users and their software users. It's easier for <redacted> clients to use the information from <redacted> and to download it themselves or have the web services themselves and so on. So they make it easier for their clients to use the open data, which is what we want.	P05: Offer process automation
And where there's still a bit more debate or the debates undecided is, OK, when governments publish open data, they're always in open formats. For instance, <redacted> is using open data that should have proprietary formats, and they are serving their own users. That's their own right but from a reuse point of view, if their client use it in their own process, well why would you care about. For instance, if it's municipality that's using <redacted> software, they use this open data that's maybe originally published original sources <redacted>, it's harvested by <redacted>, it's served in their proprietary format to municipality and municipality creates some kinds of open data service for their inhabitants based on a proprietary <redacted> format. Then the open data or and the derived open data set is a little bit less open. I mean, the data is still open, but it's not an open data format. So then suddenly people, for instance citizens if they want to interact with this data, and there's a lot of web technology now, so it's not that complex, but the problem is there are municipalities that use <redacted> software, so if you want to apply for a permit for an event, we have a tool that we provide base map and then you have to draw, OK, here is the podium, here is the first aid thing, if something go wrong the ambulance can still enter from here and from here but there will gate over here, and then suddenly people have to use a proprietary format and proprietary system to do it. Is that a problem or not, I think that's still undecided. I think the people that are most strict say no, it should always be open format because they're still in a way, if they're lucky, if you knew if it's getting too expensive, I want to go to their competitor. It's almost impossible to change vendors because it's everywhere in the processes and in a way that's may hinder -- not saying that it's actually happening -- but it may hinder the way our government works may because it's almost impossible to go to another vendor, it's becoming more expensive. Because there's no real competition on price anymore because the cost of changing is so high that you accept that basically, your monthly fees are 10% higher than with your competitor because it will cost you an enormous amount to have the whole transformation process done. Then it's something that is frowned upon within the government	P06: Develop open-source tooling
And at the same time, an intermediary like <redacted> is also really important because although they're a huge driver towards open standards and open formats and maybe even more open source tooling in the future	P06: Develop open-source tooling
At the moment, the developments in the 3D side on digital twins and so the developments are really starting to catch up, so to say. So there has to be done a lot on our side for the with the data and on their side with software and the translations and whatever. So it can always be better, but I can't mentioned specific things, let's say	P09: Foster public-private collaboration

Public-private partnerships are in my view a key mechanism for enhancing the role that data intermediaries can play in an open data ecosystem. They could be defined as long-term contracts between a government agency and a private entity with the objective of providing a public asset or service and in which the private party assumes a significant portion of the responsibility, risks and, generally, the potential benefits. While such partnerships have been used successfully in many more traditional and long-established sectors, such as large public infrastructure, it is a field that has not yet been fully explored when it comes to working with data. However, public and private entities share an interest in having high quality, accessible and cost-effective data, and that is why they are beginning to explore the new opportunities offered by these collaborative models when sharing and exploiting data, and take advantage of the potential of using private data to solve public problems, particularly in the area of smart cities.	P09: Foster public-private collaboration
So we want to know what the civil people need so the government can implement this. We want to know where a lift doesn't work so we can repair these lifts. So we want to collect this information, this data from the civil people so we can do better services for the people and not only work on: we have money and we think about where we can be implemented.	P10: Foster public-civic collaboration
Because <redacted> started as a collaboration between four partners officially, they were the four funding partners. Basically we were the 5th partner, which we acted only as a knowledge partner. They were the <redacted>. And from the beginning, because I think the first run, the creation of <redacted> as a platform and project, it was also hosted at <redacted>, so the people were working in <redacted>. After that, one of the part, the question is now we've built the platform, who's going to maintain it? And then it was decided that <redacted> was going to maintain it.	P11: Implement multistakeholder collaboration
So our role is to create the standards and also do that always in an open process so that all stakeholders are involved and are participating from the start, because it's really important to invest in participation from the beginning of the process because it influences the uptake of these standards afterwards. Because otherwise it's always "ahh it's not invented here, my standard is better, I'm already using this"	P11: Implement multistakeholder collaboration
My goal is as an <redacted>, is to improve the work of this civic society, the civic tech to have certain kind of an exchange where they can tell me on which civic tech projects they're working and which civic tech projects need to be implemented from the government themselves. From the past, there was a learning process that they have shown us what can we do and they did it themselves. And then when they stopped, then we said we can, we have to do it ourselves. But there has to be more network with the civil society that they can approach us and tell us what they're working for. Therefore, I have these hack days where I have a good exchange which ideas they're generating to profit from it and go into the government and tell them these are projects where we can put money into and what we can implement.	P13: Invest in civic tech
At the beginning these intermediaries were very important to show the profit of open data. And they're still important for the ecosystems that we are working in this in this ecosystems.	P14: Showcase open data value
Show benefits. Show countries, where it works. I think <redacted> would be a good example because they have open data, they have it nationwide, they have good quality. Still, the question is if it would be a model for <redacted>. But whatever, it would be easier for us saying all our neighbors can do it, why the <redacted> is incapable? Like in footballs, <redacted> are better than us. That's always a challenge, even to <redacted> agencies, and even <redacted> does, <redacted> does, <redacted> does? If you say all our neighbors do but only we are not able to do it, that probably would be a situation where the <redacted> would be a little bit frustrated and that would make perhaps them move. But I'm still not sure	P14: Showcase open data value
So the intermediaries can work also from the economy side or the civil side on working with open data, on returning open data to the government and also showing them what they can do with open government data	P14: Showcase open data value P15: Promote open non-govt data
On one hand, in the past that played a very important role doing pressure on politics or senators or state secretaries to say this is an important topic and that we want open data and open data is useful. So they were good intermediaries to do pressure in the government so that we can get money for open data and that we can get organizational change so we can implement this open data officers and of course show always the use of open data for the economy and also the civil community	P15: Promote open non-govt data
we also have a consultancy, the <redacted>, for example consulting the e-government, the departments and district offices because of the law. But a lot of companies from the economy asked whether there can be an open data consultancy also for small and middle sized companies and also for startups, showing them how they can work with data and why open data is very important for the ecosystems of renewable energies or the ecosystems of mobility. Because the ecosystems of mobility in <redacted> is profiting from publishing their data of lending bikes or just different kind of mobilities where they can build up a good infrastructure to see which bike service is located where, where there's an e-scooter. So there's seen the need of a central mobile platform and they wanted the state of <redacted> to accomplish a central mobile platform for all these companies. And what we need to get in return of this economy companies is that they publish data on this portal for other people who work in this closed -- how to say -- a data pool.	P15: Promote open non-govt data P16: Provide consultancy
By using information from other organizations, the quality of this information increases as the organization responsible for the data is informed of errors and improvements-- the feedback of the information	P16: Provide consultancy

And because -- we work with our colleagues in many ways. At the <inaudible> they need. If they are GIS -- they know how to deal with ArcGIS or so, they do everything. If they don't know nothing, they give us the data, we find the -- if it's structure or not, we talk to them and we clean them and position them in a map. And between these two positions, whichever the point you are we want, we give you the tools to deliver data to citizens and of course to our own <redacted>	P16: Provide consultancy
we have been running this geoportal only <redacted> years. And in <redacted> years we became the main information system within <redacted> because this point of view, that everything is everything happens somewhere, as we used to say, is the main idea that we deliver to our colleagues. So we want to spread it through <redacted>. I think it's a very big idea and a good one because of what we are – sorry I didn't find the word – our colleagues have been giving feedback to us. So they say that running their own businesses have changed since they started to use geographical information system because -- they had numbers, they had a total amount, they have distributions in boroughs, for example, it's very clear when you see an image of what you are spending money in and how it's related to some boroughs, and some other boroughs are not given the same amount of money, maybe they are suffering from worse city conditions. So you have to put more money in them instead of the other ones. So it's another way to keep vulnerability away and to equal the efforts in the different boroughs of the city.	P16: Provide consultancy P18: Facilitate internal reuse of open data
We have several groups in that we share information and share ideas and what we think. One of the forums is the <redacted> that take place here in <redacted> every year. And we share information with many, not only <redacted>, but his partners and other administrations. We have another event that is intended for a local and state administrations and we share information with our fellows also from <redacted>. And there are cartographic events that take place every four years in a different city in <redacted>. Well, yeah, as soon as the there is an event to share information there we go.	P17: Streamline cross-administrative processes
I think it's important to compare this to the current state of INSPIRE, because INSPIRE was always meant to be an open data ecosystem and that has been designed with specific purpose in mind, [e.g.] sharing environmental data in case of some kind of disaster, and the idea was disasters never stopped at administrative boundaries, so it should be -- you should be able to combine data across borders and it need to be available because as soon as the disaster hit you don't have time to start thinking about the creations of data or publication of data, it should be made available already. So in the case of emergency you can just access the data immediately.	P17: Streamline cross-administrative processes
And on the other hand, this accessible and free information is of great value for other administrations and for companies of the sector. It allows them to create value added data and services from the basic information which they obtain for free with the cost savings that it entails	P18: Facilitate internal reuse of open data
I think the last thing is making data available in a way that it's more user friendly. Because still the biggest user of governmental open data are other governments, other governmental bodies. So in a way it's quite weird in the way of thinking, the government was not allowed to publish data in such a way that another government was not able to use it – to say no, if another government wants to use my data, you should hire a private company to adjust it, make more flexible or have a little twist. It's just weird. So I'm really hoping that in this current movement into data space, or federated architecture, whatever you want to call them, there's much more collaboration. And it's an understanding that private parties need the government to sometimes say, OK, this data, this is most important data. For instance, if you are applying for building permits or do your analysis based on this data source, this data source is the data source. We will guarantee you that -- we will not say you're using the wrong data and as a result, you're not getting your permit. I think we're really creating a more level playing field for all actors and they're, I think public and private parties, really need each other.	P18: Facilitate internal reuse of open data
But it allows the geographic data to reach a border audience, and that also means that we receive more feedback from our users. That brings us a quality control of data because every day people contact us to say, I see this error or that other. It also allows us to direct our strategy towards what citizens and other public administrations and even companies are demanding. So it's like a contact between the public and us.	P19: Facilitate feedback on open data
I think that the data intermediaries could play this bridging role, that bridge the actual user with the data providers. That could be a very nice role for them to also take	P19: Facilitate feedback on open data
we also have our platform <redacted> where we also proactively ask the civil society to participate in these questionnaire, in this survey, so we can find out what is important for them, which data is important for them, which projects they want to be included	P19: Facilitate feedback on open data
And there you say, OK, of course I want to access, for instance, the building registry, but I want to combine the data with another source because I have this question a million times a day. Those kinds of user request always pop up on a service provider and not on the individual data provider. I'm only concerned with addresses, so combining it with other data, yeah, don't ask me. So basically this intermediaries playing a really important role in the transformation being the supply driven infrastructure into a much more demand driven infrastructure. Because now users have an entry point where they can post on request where they can say OK this is not nice format but it would really be helpful if the data is available in that format as well. And when you're serving it for hundreds of data sets it is much easier to see those kinds of trends and to understand what kind of formats you should invest. So I think these intermediaries like <redacted> are really important.	P19: Facilitate feedback on open data